



# Clasificación de manzanas utilizando visión artificial y redes neuronales artificiales

## Classification of apples using artificial vision and artificial neural networks

Mota-Delfin, C<sup>1</sup>., Juárez-González, C<sup>2</sup> y Olgún-Rojas, J. C<sup>3</sup>

### Resumen

El valor agregado en un fruto puede incrementarse con un buen manejo postcosecha. La clasificación en diferentes parámetros es de las operaciones más importantes. En las pequeñas empresas se realiza de forma manual obteniendo deficiencias en la calidad del producto. Estos problemas se podrían resolver o disminuir con la implementación de algoritmos inteligentes que en este caso incluyen visión artificial y redes neuronales artificiales. En este proyecto se presenta la clasificación de manzanas mediante un algoritmo inteligente, usando una red neuronal convolucional (CNN), la cual se desarrolla utilizando librerías de código abierto (OpenCV, Tensorflow y Keras) en Python con una estructura de diferentes capas convolucionales y MaxPooling, para un conjunto de 2,800 imágenes de 128x128 píxeles, de las cuales el 80% se utilizaron para entrenamiento y el 20% para hacer una prueba de la red, obteniendo una precisión del 98.3% y del 95.36%, respectivamente. Después del entrenamiento, se realizó una clasificación con un video en tiempo real, obteniendo una precisión del 92.25%. Asimismo, se explora la posibilidad de la utilización en la industria con la clasificación por otras características visuales del fruto como tamaño, color, forma, etc.

**Palabras clave:** algoritmo inteligente; convolución; entrenamiento; prueba; clasificación

### Abstract

The added value in a fruit can be increased with a good postharvest handling. The classification in different parameters is one of the most important operations. In small companies it is done manually, obtaining deficiencies in the quality of the product. These problems could be solved or reduced with the implementation of intelligent algorithms that in this case include artificial vision and artificial neural networks. In this project is presented the classification of apples through an intelligent algorithm, using a convolutional neural network (CNN), which is developed using Open Source libraries (OpenCV, Tensorflow and Keras) in Python with a structure of different convolutional layers and MaxPooling, for a dataset of 2,800 images of 128x128 pixels, of which 80% were used for training and 20% for test of the network, obtaining an accuracy of 98.3% and 95.36%, respectively. After the training a classification was made with a video in real time, obtaining an accuracy of 92.25%. Likewise, the possibility of using it in the industry is explored with the classification by other visual characteristics of the fruit such as size, color, shape, etc.

**Keywords:** intelligent algorithm; convolution; training; test; classification

<sup>1</sup> Estudiante de Ingeniería Mecánica Agrícola. Universidad Autónoma Chapingo. Chapingo-México. Email: [kincanekmd@gmail.com](mailto:kincanekmd@gmail.com)

<sup>2</sup> Estudiante de Ingeniería Mecánica Agrícola. Universidad Autónoma Chapingo. Chapingo-México. Email: [carlos.juarez.gonzalez904@gmail.com](mailto:carlos.juarez.gonzalez904@gmail.com)

<sup>3</sup> Profesor del Departamento de Ingeniería Mecánica Agrícola de la Universidad Autónoma Chapingo. Chapingo-México

## 1. Introducción

El buen manejo post cosecha de un fruto siempre es fundamental para asegurar su comercialización. La clasificación de este ayuda a mantener una mejor apariencia y sirve para separarlas en lo que generalmente se comercializa, como cajas, las que pueden contener frutos del mismo tipo y tamaño. Las diferentes necesidades del mercado de frutas hacen que la clasificación de estas sea de suma importancia y las soluciones comunes son hacerlo de manera manual, con personas que se especializan en este trabajo, sin embargo, el rendimiento manual no es suficiente ni rentable. Para evitar las complicaciones y retrasos en línea de producción que se generan al llevar a cabo la clasificación manual, se suele utilizar máquinas clasificadoras que realizan el proceso automáticamente (Quiminet, 2013), que debido a la complejidad de sus mecanismos adquieren un costo elevado para productores de pequeña escala que aún buscan un mercado establecido.

Si bien, hemos de valorar la mejor calidad del trabajo manual de clasificación cuando se trata de características cualitativas y la de las máquinas cuando son características cuantitativas, por ello en el presente trabajo se trata de conjuntar estas dos ventajas con el uso de algoritmos inteligentes que comprenden visión artificial y redes neuronales artificiales para la posible solución del problema de clasificación de frutos. Para iniciar este trabajo, elegimos la manzana ya que, en el mundo se conocen 7,500 variedades y en Estados Unidos 2,500 aunque no todas son comerciales (ABC, s.f.) y sus características más importantes como color, forma y tamaño diferencian unas de otras, algunas de las cuales son difíciles de diferenciar entre variedades.

El objetivo de este trabajo es desarrollar un algoritmo inteligente capaz de diferenciar las principales características de cuatro variedades de manzanas y clasificarlas, reunir los datos necesarios para su entrenamiento y evaluar su rendimiento con la puesta en marcha en tiempo real, así como comprender su utilidad y facilidad de utilización para cualquier otro fruto con otras características distintas.

## 2. Materiales y métodos

El desarrollo de este trabajo se llevó a cabo en las instalaciones del Departamento de Ingeniería Mecánica Agrícola de la Universidad Autónoma Chapingo, ubicada en km 38.5 Carretera México-Texcoco, Chapingo, Estado de México. La metodología utilizada consistió de los siguientes pasos:

### 2.1 Obtención de datos

Para la obtención de los datos se utilizó una metodología similar a la de (Horea & Mihai, s.f.), para este caso, se utilizaron 8 manzanas de 4 variedades diferentes (Tabla 1), a cada una se le colocó un palillo de madera forrado de color blanco lo suficientemente largo para poder sostener la fruta en frente de un fondo blanco y grabar un video de alrededor de 15 segundos mientras giramos la manzana en tres posiciones diferentes, utilizando una cámara Logitech Pro c920 y el programa *Logitech Webcam Software*.

Posteriormente, haciendo uso del software *Free Video to JPG Converter* se extrajeron 700 fotogramas del contenido de los videos capturados de cada variedad de manzana. Por lo que el conjunto de datos lo conformaron 2,800 imágenes, de la cuales se utilizó el 20% para la prueba de la red y el 80% para el entrenamiento.

**Tabla 1.** Variedad y numero de imagen de cada manzana

<b>Variedad de manzana</b>	<b>Imágenes para entrenamiento</b>	<b>Imágenes para Prueba</b>
<b>Gala</b>	560	140
<b>Gold</b>	560	140
<b>Granny Smith</b>	560	140
<b>Red Delicious</b>	560	140

Aún con el palillo disfrazado del mismo color del fondo, este sigue notándose, por lo que se desarrolló un algoritmo para extraer la imagen del fondo, que consiste en determinar el color de las manzanas aplicando filtros de colores y así ubicar el centroide en la imagen, partiendo de este, generar un cuadro para recortarla y guardarla. Debido a que hay cambios de colores en las manzanas el código se modificaba continuamente de acuerdo a consideración de los resultados. Posteriormente el algoritmo escala la imagen al tamaño de 128 x128 pixeles, en otros casos se han utilizado tamaños de 28 x 28 como en la base de datos de MNIST (LeCun, Y., Cortés, C., & JC Burges, C., s.f.), por lo que una primera capa de red neuronal ordinaria tendría  $28 \times 28 = 784$  pesos, la cual parece manejable, pero es evidente que esta estructura no funciona bien con imágenes más grandes y con una imagen de mayor tamaño a 200 x 200 pixeles sería un desperdicio y conduciría rápidamente al sobreajuste (López Briega, 2016).

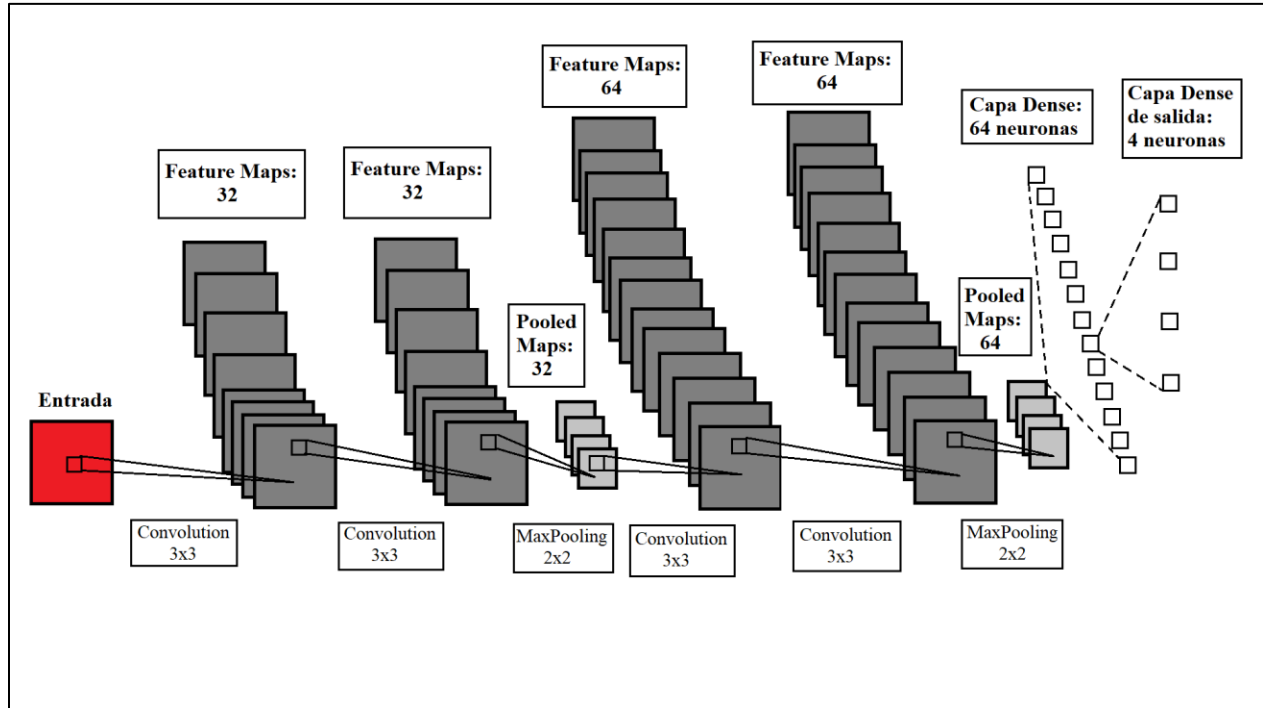
## **2.2 Estructura de la red neuronal utilizada**

Para este proyecto se utilizó un modelo secuencial en el cual se emplea una red neuronal convolucional que es un tipo de red neuronal artificial, donde las neuronas corresponden a campos receptivos de una manera muy similar a las neuronas en la corteza visual primaria (V1) de un cerebro biológico (LeCun, Y., Cortés, C., & JC Burges, C., s.f.). Es esta utilizamos capas convolucionales seguidas de su función de activación tipo ReLU (Rectified Linear Unit) así como capas MaxPooling, capas totalmente conectadas y capas de pérdida. La tabla 2 muestra el orden de capas, y en la figura 1 se representa su arquitectura.

**Tabla 2.** Estructura de la red neuronal utilizada en este trabajo

<b>Tipo De Capa</b>	<b>Dimensiones</b>	<b>Salidas</b>
<b>Convolutional</b>	3x3	32
<b>Convolutional</b>	3x3	32
<b>MaxPooling</b>	2x2	-
<b>Convolutional</b>	3x3	64
<b>Convolutional</b>	3x3	64
<b>MaxPooling</b>	2x2	-
<b>Fully connected</b>	64	64
<b>SoftMax</b>	64	4

### 2.3 Evaluación del modelo

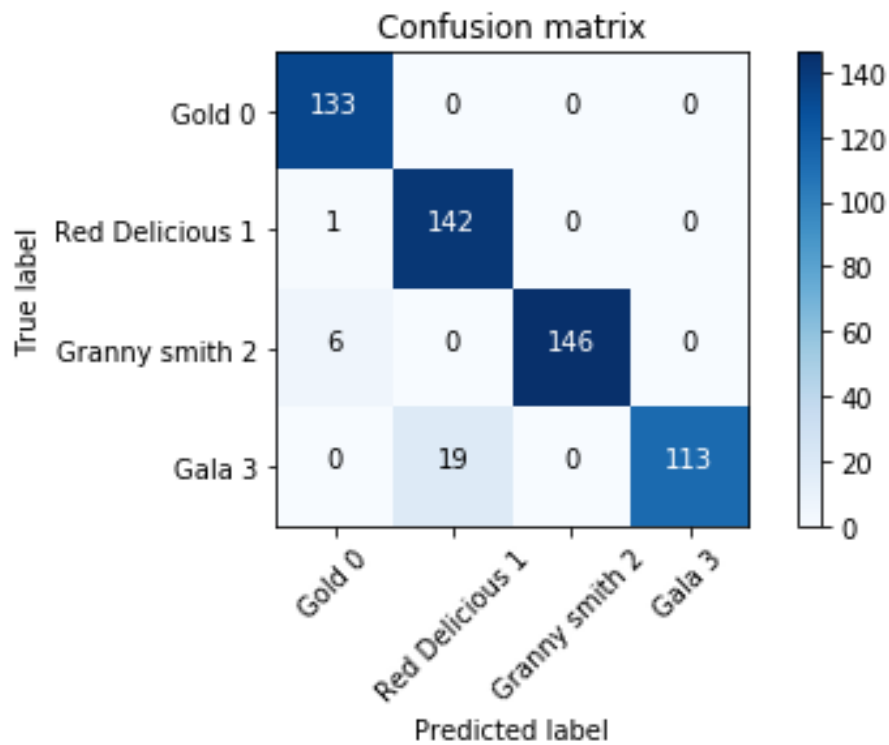


**Figura 1.** Arquitectura de la red neuronal utilizada en este trabajo

Para medir el rendimiento del algoritmo de aprendizaje automático supervisado se hizo uso de una matriz de confusión en la cual cada fila representa las instancias de una clase real y cada columna representa las instancias de una clase predicha, en la evaluación de la red neuronal. Seguidamente se procesó; un conjunto de datos obtenidos de Kaggle Inc (2018) de manzanas Grand Smith y Red Delicious; 25 fotografías de manzanas Gold; y un video en tiempo real con manzanas Gala, que añadiendo un código que lee todos los fotogramas del video, se corroboró el rendimiento de predicción de las manzanas.

### 3. Resultados y discusión

El entrenamiento mostró una precisión del 98.3%, sin embargo, al medir el rendimiento del algoritmo con las imágenes de prueba seleccionadas aleatoriamente, se clasificaron erróneamente 26 de las 560 imágenes, (las cuales se muestran en la figura 2), por lo que muestra una precisión del 95.36%.



**Figura 2.** Matriz de confusión de la prueba de la red neuronal

Esta prueba de rendimiento no asegura que el algoritmo se desempeñe con la misma precisión al ser sometido a una clasificación en tiempo real y con datos diferentes a los utilizados para el entrenamiento y evaluación de la red neuronal.

Los resultados de la evaluación de predicción del modelo muestran una precisión del 92.25% y se presentan en la tabla 3.

**Tabla 3.** Resultados de la prueba con el modelo entrenado y evaluado

		VARIEDAD PREDICHA POR EL MODELO			
		Gala	Gold	Granny Smith	Red Delicious
VARIEDAD REAL	Gala	386	0	5	24
	Gold	0	24	1	0
	Granny Smith	0	0	152	12
	Red Delicious	0	0	16	128

#### 4. Conclusiones

De acuerdo a los resultados obtenidos, la aplicación de este método en la clasificación de manzanas da muy buenos rendimientos, los cuales se pueden mejorar ampliando nuestra base de datos para el entrenamiento en cuanto en calidad de imagen y calidad de la iluminación, ya que estos aspectos generan ruido que afectan la precisión de la red.

También es importante mencionar que la prueba del modelo se realizó en un ambiente con mucho ruido, en cuanto al fondo de la imagen e iluminación, lo cual, en la aplicación se realizará con un ambiente controlado, como en una línea de proceso, pudiendo a si mejorar su precisión.

La elaboración de este proyecto da pauta para la continuación de clasificación no solo de las manzanas en clases, si no también considerar aspectos como son; enfermedades en la manzana, golpes, tamaños, etc. aspectos visibles o no visibles al ojo humano utilizando camaras multiespectrales.

#### 5. Referencias bibliográficas

ABC DE FRUTAS Y VERDURAS. (s/f). *MANUAL TÉCNICO DE FRUTAS Y VERDURAS*. Morelos: Grupo PM.

Horea, M., & Mihai, O. (s.f.). Fruit recognition from images using deep learning.

Kaggle Inc. (2018). *Fruits 360 dataset*. Obtenido de <https://www.kaggle.com/moltean/fruits/data> (08/08/2018).

LeCun, Y., Cortés, C., & JC Burges, C. (s.f.). *LA BASE DE DATOS DE MNIST*. Obtenido de <http://yann.lecun.com/exdb/mnist/> (06/08/2018).

López Briega, R. (02 de Agosto de 2016). *Redes neuronales convolucionales con TensorFlow*. Obtenido de <https://relopezbriega.github.io/blog/2016/08/02/redes-neuronales-convolucionales-con-tensorflow/> (06/08/18).

Quiminet. (Marzo de 2013). *La manera más eficaz de clasificar frutas y verduras*. Obtenido de <https://www.quiminet.com/articulos/la-manera-mas-eficaz-de-clasificar-frutas-y-verduras-3457632.htm>. (04/08/18)